



FILE FORMATS BEST PRACTICE FOR ALL PUBLIC AGENCIES

Date Effective: December 1, 2007

Approved: November 28, 2007 (Last Revised: January 8, 2010)

PURPOSE

The purpose of this best practice is to establish a set of statewide recommendations for converting records to file formats that support interoperability and/or the long-term preservation of public records. This best practice is not intended to be a statement of the current ability of public agencies to convert records. It is a statement of goals and expectations. The realization of such goals and expectations will result in more effective records management.

SCOPE

This best practice applies to all records created or received by public agencies.

STATEMENT OF AUTHORITY

Pursuant to 1 V.S.A. § 317, 3 V.S.A. § 117, 3 V.S.A. § 218, 3 V.S.A. § 2222, and 3 V.S.A. § 2283b the Vermont State Archives and Records Administration (Office of the Secretary of State) and the Department of Information and Innovation (Agency of Administration) are authorized to establish and promulgate standards, procedures and techniques for the effective management of public records.

STATEMENT OF BENEFITS

There are over 1,000 file formats available for use by public agencies. Some formats are more stable than others relative to long-term preservation. It is important to balance the practical use of a file format with its ability to be opened, accessed, and read in both the short- and long-term. Public agencies will benefit from creating records in file formats that are amenable to long-term access and future migrations and conversions.

STATEMENT OF RESPONSIBILITY

Maintaining complete and trustworthy records over time is a shared responsibility. Establishing and operating effective recordkeeping systems and practices requires a multidisciplinary approach. Public agencies should make effective use of the necessary range of expertise available throughout the State of Vermont. This includes expertise in archives, records and information management, information technology, business process management, risk management, and law.

CONTACT

Questions about this best practice may be directed to the Vermont State Archivist or Chief Information Officer.



DEFINITIONS

Extensible: The ability to be expanded or customized by changing or adding new control structures, statements, or data types.

File Format: The layout/structure of information or data within a record (file). For electronic records, the file format is generally identified by the record's extension (i.e. *.txt*, *.doc*, *.pdf*, *.exe*, etc.). Most electronic file formats require specific software applications to open, access, and read the records.

File Format Specification: Specific information concerning how information or data within a file is arranged, subdivided, sized, encoded, sequenced, related to other data, etc. within a certain file format that permits the file to be properly opened, interpreted, and represented for use and inspection. Specifications are predefined structural and organizational components of a particular file format.

Interoperability: The ability of different information technology systems and software applications to: communicate; exchange data accurately, effectively, and consistently; and use the information that has been exchanged.

Non-Proprietary: The specifications for hardware and/or software are available for independent review, development, and distribution.

Preservation: The act of keeping a record in readable, accessible, and unaltered condition; to maintain unchanged and intact.

Public Agency: Any agency, board, department, commission, committee, branch, instrumentality, or authority of the state or any agency, board, committee, department, branch, instrumentality, commission, or authority of any political subdivision of the state (1 V.S.A. § 317a).

Records: All papers, documents, machine readable materials, or any other written or recorded matters, regardless of their physical form or characteristics, that are produced or acquired in the course of agency business (1 V.S.A. § 317b).

Source Code: Programming statements and instructions written by a human in plain text (i.e. human readable alphanumeric characters).

SPECIFIC GUIDELINES

For determining specific file formats that meet the recommendations of this best practice, the Information Strategies: Taskforce on Archives, Records, and Technology (iSTART) has issued a *File Formats Guideline*.

BEST PRACTICE REVIEW

The State Archives and Department of Information and Innovation will review this best practice as necessary in consultation with public agencies, research communities, and other interested parties.



FILE FORMATS BEST PRACTICE

- 1) File format should be publicly and openly documented.**
 - Format specifications are published and available for use with minimal legal restrictions.
- 2) File format should be non-proprietary.**
 - Format is supported by more than one developer and can be accessed with different software systems.
- 3) File format should be widely applied.**
 - Format is commonly used by the public and private sector.
- 4) File format should be self-documenting.**
 - Format is not dependent on external documentation to determine the content, context, and/or structure of the record.
 - Source code is human-readable.
- 5) File format should be extensible.**
 - Format can be expanded and modified by users and developers.
- 6) File format should be opened, read, and accessed using readily available tools.**
 - Format is recognized by standard office tools and applications and human-readable representations are consistent with the underlying data representations.
- 7) File format should be easily migrated to other platforms or storage media.**
 - The original structure and content of the record should remain unchanged when the record is migrated from one platform or storage medium to another.
 - If the new platform or storage medium cannot support a record's existing file format, the record shall be converted to a new file format that is supported.
- 8) File format shall be easily converted to other file formats.**
 - The content and structure of the record, based on the record's native format, should be retained in the new format to the extent necessary to ensure the trustworthiness and completeness of a record upon conversion.
 - Relationships of data in the record (i.e. spreadsheet formulas, database fields, hyperlinks, etc.) should be retained when the record is converted from one file format to another.
 - Loss of data or structure should be minimized at all times. Loss to either may affect legal requirements for complete and authentic records.



9) The selection of a file format, and any decisions to migrate or convert files, should be the result of careful planning.

- File format should meet all legal and operational requirements.
- The risks and advantages of using a file format should be identified and documented prior to making a selection.
- The scope to which a record's content or structure may be loss or altered should be known prior to migrating records from one platform or storage media to another.
- The scope to which a record's content or structure may be loss or altered should be known and documented prior to converting records from one file format to another.

REVISION HISTORY

2010-01-08	Removed reference to Commissioner of Buildings and General Services under "Contact" that was missed in 07-31-2008 revision.
2008-10-17	Replaced the word "policy" with "best practice" to reflect similar updates that have been made to the <i>Records Management Best Practice</i> .
2008-07-31	Removed references to the Department of Buildings and General Services and updated statutory authorities pursuant to Acts 96 and 110 of 2008, which became effective July 1, 2008. Replaced "Vermont State Archives" with "Vermont State Archives and Records Administration."
2008-01-30	Statement of purpose was expanded to articulate that the policy is not representative of the current abilities of public agencies. Policy statements were revised to use the word "should" instead of "shall." Definition for "source code" was added to the Definitions section, and iSTART replaced the individual agencies as the body that issued the <i>File Formats Guidelines</i> .